INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
20–1584/2021/25–4–768–776
DOI: 10.17957/IJAB/15.1728
http://www.fspublishers.org

Full Length Article

DNA Fingerprinting and Genetic Diversity Assessment of GM Cotton Genotypes for Protection of Plant Breeder Rights

Shakra Jamil, Rahil Shahzad*, Muhammad Zaffar Iqbal, Erum Yasmeen and Sajid Ur Rahman
Agricultural Biotechnology Research Institute, Ayub Agricultural Research Institute, Faisalabad, Punjab Pakistan
*For Correspondence: rahilshahzad91@gmail.com
Received: 3 November 2020; Accepted: 28 December 2020; Published 25 March 2021

Abstract

DNA fingerprinting is rapid, easy, and efficient method for discrimination, identification and characterization of various genotypes for protection of plant breeder's rights (PBRs). Present study was designed for DNA fingerprinting and genetic diversity assessment of 25 GM cotton genotypes (possessing *Cry1Ac gene*) using 297 SSR markers through conventional PCR and Polyacrylamide gel electrophoresis. Out of 297 SSR markers, 25 markers were not amplified, 28 were monomorphic and 244 were polymorphic. A total of 1537 alleles were amplified among which 1294 (84.18%) were polymorphic. PIC value in our study ranged from 0.08 to 0.93 with an average of 0.73. Unique allelic pattern was observed for nineteen genotypes whereas six genotypes were identified using two-step identification methods. The UPGMA dendrogram divided the genotypes into two distinct clusters. Cluster I was comprised of 20 genotypes whereas cluster II was comprised of four genotypes. MNH-1020 did not obey any clustering and remained separated. The results of the structure analysis were complementary to cluster analysis and the population was divided into two subgroups. Our results evidenced narrow genetic base of the cotton genotypes cultivated in Punjab Pakistan due to use of common parents in the pedigree/parentage. Further, we proposed a core set of markers for future DNA fingerprinting and genetic diversity studies. The information generated in this study will be helpful in variety registration and subsequent protection under PBRs. Further our findings will be useful in selection of SSR markers for future studies which are focused on DNA fingerprinting and genetic diversity assessment.

Keywords: Cluster analysis; Plant breeder rights; Polymorphic information content; Structure analysis

Introduction

Cotton (*Gossypium spp.*) also known as "White gold" is one of the major cash crops around the globe which is mainly cultivated to produce raw fiber for the textile industry (Singh 2017; Rehman *et al.* 2019; Jans *et al.* 2020). World fiber production equaled approximately 110 million metric tons in 2018, including 32 million tons of natural fibers and 79 million tons of chemical fibers. Cotton accounted for 80% of natural fiber production by weight (Townsend 2020) which shows its significance in international economies. Pakistan is the fourth largest lint producer of cotton (Shuli *et al.* 2018; Lalwani 2020).

Distinctness, uniformity, and stability (DUS) testing remain the sole scientific criteria for the protection and registration of new varieties in past (Pourabed *et al.* 2015). Earlier morphological and biochemical markers were used for DUS testing. The use of these markers produces inconsistent results because morphological and biochemical markers are influenced by the plant age, the environment and other factors. With the availability of molecular markers, it became possible to conduct rapid and accurate identification at the DNA level without the impact of environmental factors (Iqbal *et al.* 2017; Santhy *et al.* 2019). DNA fingerprinting is the rapid, easy, and most common method to discriminate, identify and characterize various cultivars to protect PBRs and promote marker-assisted breeding (Kalia *et al.* 2011). The technique has been revolutionized since the past three decades to distinguish the DNA polymorphism, biological identification, and documentation of species. Genetic profiling recapitulates the biological determination of species as well as traceability of diverse crop samples using the short tandem repeats. Through this PCR based approach, individual plant hybrids/varieties can be identified by acquiring a specific pattern of genetic profiles (Zhang *et al.* 2013).

The DNA fingerprints are stored in databases and sequences could be used for direct selection and identification of cotton hybrids and parents for future crop production programs. Moreover, the International Union for the Protection of new Varieties of Plants (UPOV) has encouraged the use of molecular markers in DUS testing for

the identification of crop cultivars (UPOV-BMT 2002). Molecular markers are frequently used for effective selection, robust assessment of polymorphism and to explore the relativity of diverse genetic groups of cultivars with their wild relatives (Shah *et al.* 2009; Király *et al.* 2012). Previously, a restricted cotton gene pool has been classified by using Random amplified polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP). SSR/Microsatellites are proven to be an ideal tool for DUS testing of new varieties because of high polymorphism, multi-allelic, co-dominant inheritance, good reproducibility, abundant distribution all over the genome and short amplification product and widely used for molecular characterization of genotypes to accelerate the effective selection (Jamil *et al.* 2020). About >1000 primers are identified from the cotton genome that is available in genome libraries (Nguyen *et al.* 2004; Yu *et al.* 2014).

Although some studies were conducted for DNA fingerprinting and genetic diversity assessment of cotton varieties in Pakistan previously (Mumtaz *et al.* 2010; Ullah *et al.* 2012). As far as Mumtaz *et al.* (2010) is concerned they have used RAPDs markers which are less reliable and non-reproducible. Ullah *et al.* (2012) although used SSR markers for DNA fingerprinting however genotypes used in their study were all primitive and number of SSR markers (104) used were relatively low which are unable to reveal genetic diversity in Pakistani cotton genotypes having narrow genetic base. Keeping in view of above said facts in our study we utilized 297 SSR markers for DNA fingerprinting and genetic diversity analysis of 25 cotton genotypes. Cluster analysis was conducted for estimation of genetic distance and to provide a reliable picture of a diverse grouping of genotypes for effective utilization of genetic information in cotton breeding programs. Structure analysis and dendrogram provides an insight into different sets of allelic richness in GM cotton genotypes. DNA fingerprints of GM cotton will provide a molecular basis to identify and authenticate the seed purity in the market.

Materials and Methods

Plant material

The research work was conducted at Agricultural Biotechnology Research Institute, Ayub Agricultural Research Institute Faisalabad, Pakistan for DNA fingerprinting of cotton genotypes for variety protection and registration under Plant Breeders Rights Rules 2017. Seeds of 25 GM cotton genotypes (Pure-lines) were obtained from various institutes and stations from all over Punjab (Table 1) and were sown in pots in the greenhouse at 28°C following the standard agriculture practices. Each genotype was planted in 5 different pots wherein, each pot contained 2 seeds per genotype. After germination and seedling development till 3–4 leaves, 05 seedlings for each genotype were harvested and stored at -40°C for DNA extraction.

DNA isolation and PCR amplification

DNA was isolated from 100 mg of young leaves of GM cotton plants. The leaves were finely ground to powder using liquid nitrogen while DNA extraction was performed using the modified cetyltrimethylammonium bromide (CTAB) method (Allen *et al.* 2006). DNA samples for all genotypes were analyzed for the quality and quantity by NanoDrop spectrophotometer and also running them on Agarose Gel Electrophoresis as elaborated in our previous study (Jamil *et al.* 2020).

PCR was assembled using 2X DreamTaq Green PCR master mix ThermoFisher Scientific (K1082) as recommended by the manufacturer. The master mix aids us in direct loading the samples (PCR product) on gel and green dye does not cause any inconvenience during PCR reaction. For 2X we prepared 50 μL PCR reaction mixture which was comprised of 25 μL master mix, 200 ng template DNA, 2 μM primer (forward & reverse) and volume make up to 50 μL using nuclease-free water. PCR profile was set as follows: 1 cycle of initial denaturation at 95°C for 5 min, 35 cycles of denaturation at 95°C for 60 sec, annealing at 55°C for 1 min and extension at 72°C for 1 min, the final extension for 5 min at 72°C. The PCR product was stored at 4°C before electrophoresis.

Poly acrylamide gel electrophoresis (PAGE)

PCR amplicons were analyzed on Electrophoresis System model POWERPRO-3AMP (cleaver scientific limited) using 6% (W/V) polyacrylamide gel and performed at 16 watts power. PAGE gel was stained by Silver nitrate staining for visualization according to the previously described staining protocol by (Caetano-Anolles 1997). Resulting patterns were analyzed and captured using Syngene trans-illuminator Gel Documentation System.

Statistical analysis

The size of the PCR amplicon for each genotype was estimated by comparing them with banding patterns of 50 bp DNA ladder (Invitrogen™) loaded in the PAGE gel and scored as 1 or 0 indicating the presence and absence of particular band size. NTSYSpc 2.0 software (Rohlf 1998) and Structure v. 2.3.4 were used for statistical analysis. Structure analysis utilizes a model-based Bayesian clustering approach to obtain pedigree information that provides different sets of allelic richness in cotton (Pritchard *et al.* 2000). NTSYSpc 2.0 uses an un-weighted pair Group Method of Arithmetic Means (UPGMA) to analyze multivariate diverse data and generates a dendrogram. The structure analyses were performed according to the following parameters: No admission model; K-value (1–10); Burn-in periods: 10,000; 2 number of in-iteration burns 10,000; Markov Chain Monte Carlo Simulation: 100,000. The structure was determined by using LnP (K) values against ΔK values using Evanno Test (Evanno *et al.* 2005).

Results

SSR Polymorphism

A total of 297 SSR markers evenly distributed on 26 cotton chromosomes were used for DNA fingerprinting and genetic diversity studies of 25 cotton genotypes (Table S1). Among 297 SSR markers, 25 markers were not amplified whereas 28 were monomorphic and the remaining 244 were polymorphic. The polymorphic 244 SSR markers amplified a total of 1537 alleles among which 1294 (84.18%) were polymorphic and 243 alleles (15.82%) were monomorphic. Minimum numbers of alleles 2 were amplified by 13 SSR markers namely BNL0347, BNL2570, BNL3103, BNL3140, CIR0208, CIR0210, DPL0058, DPL0156, DPL0163, DPL0273, JESPR85, MUCS0515 and TMB2920. Maximum numbers of alleles (19) were amplified by SSR marker BNL0137 among which 16 were polymorphic (Fig. 1). Maximum polymorphic alleles (PA) 18 were amplified by BNL-228. Lowest PIC value (0.08) was observed for DPL0156 whereas the highest PIC value (0.93) was recorded for seven SSR markers *i.e*., BNL0137, BNL0387, BNL3977, JESPR220, JESPR222, MGHES44 and TMB0471 collectively. The average number of alleles and polymorphic alleles was 6.3 and 5.3 respectively. The average PIC value was 0.73 whereas the size of amplicon ranged from 80 to 1000 bp (Table 2).

DNA fingerprinting

Fifty-seven SSR markers were able to differentiate 25 cotton genotypes. There were two groups of genotypes concerning to DNA fingerprinting. Group, I comprised 19 genotypes that amplified unique alleles and were identifiable using single SSR marker. Group II was comprised of 06 genotypes which were not identifiable using unique alleles hence a two-step identification method was used for their identification (Table 3).

SSR marker BNL0119 amplified unique alleles for four genotypes *i.e*., MNH-1016, MNH-1020, BH-221, and IUB-13. Similarly, BNL0228 amplified unique allelic patterns for three genotypes *i.e*., MNH-886, FH-142, and IUB-13. IUB-13 was identifiable with the help of nine SSR markers, MNH-1016 was identifiable using six SSR markers, NIAB-878 amplified unique allelic pattern with five SSR markers, MNH-1020 and RH-668 with four markers, FH-326 and BH-221 with three markers, MNH-886, VH-327, RH-647, RH-662, SLH-8, SLH-19 and BS-15 with two markers and VH-Gulzar, FH-142, SLH-06 and BH-201 by one marker as given in Table 3.

Genetic diversity studies

The data of 244 polymorphic SSR markers were used to generate a UPGMA dendrogram to study the extent of genetic diversity among 25 cotton genotypes using the SHAN similarity matrix. The similarity coefficient between 25 cotton genotypes varied from 0.63 to 0.91. The dendrogram divided the genotypes into two distinct clusters (Fig. 2). The highest similarity was observed between FH-152 and FH-142 in cluster I sharing 91% of the genetic loci; whereas the lowest genetic similarity was observed between MNH-1020 and the rest of 24 genotypes sharing 63% of genetic loci in common. A domestic relationship exists between cultivar distribution and agro-ecological zones as is evident from the UPGMA dendrogram. Genotypes bred in different agro-ecological zones *i.e*., Multan, Sahiwal, Vehari, Faisalabad, and Bahawalpur tend to appear in the same clade in cluster analysis. However slight variation was observed for MNH-886, MNH-1020, RH-668, and BH-178 which did not follow geographical distribution (Fig. 2).

In most cases, cluster analysis results fitted well with pedigree parentage information. Genotypes lying in clade IA have a common parentage with one another except VH-383 and VH-Gulzar. Similarly, varieties present in clade IB *i.e*., except SLH-08, SLH-19, and RH-668 have one parent in common with each other. Genotypes present in clade III *i.e*., BH-201 and BH-221 do not have common parentage; whereas genotypes present in cluster II *i.e*., IUB-13 and BS-15 have shared parentage except for NIAB-878 and MNH-886 (Fig. 2 and Table 1).

Population structure studies

The model-based Bayesian approach was used to infer the population structure of twenty-five genotypes of cotton. Population Structure was determined by plotting the LnP (K) against ΔK values through Evanno Test (Evanno *et al.* 2005). The highest ΔK (10212.75) was observed for K=2 suggesting the existence of two sub-populations *i.e*., population 1 (P1) and population 2 (P2). P1 was comprised of 21 genotypes whereas P2 was comprised of four genotypes. The expected heterozygosity was high among the individuals of P1 (0.132) as compared to P2 (0.10). Whereas genetic diversity among the individuals of P2 (Fst_2 = 0.82) was high as compared to P1 (Fst_1 = 0.69) (Fig. 3). The results suggested that all genotypes originated from the Government of Punjab Agriculture set up (Ayub Agricultural Research Institute, AARI) have similar blood except MNH-886. Similarly, varieties bred from Institute other than AARI *i.e*., IUB-13, BS-15, and NIAB-878 have similar blood.

Discussion

DNA fingerprinting and genetic diversity studies are of prime importance for germplasm maintenance, PBRs protection, and seed production in cotton (Santhy *et al.* 2019). For a cotton breeder, presence of genetic variability guides for interspecific or intraspecific hybridization (Sheidai *et al.* 2014). Estimation of genetic diversity and DNA fingerprinting characterizes the individuals and assign

Table 1: List of genotypes used in the study along with pedigree/parentage information

Institute Name	Variety name	Pedigree/Parentage
CRI, Multan	MNH-1016	MNH-786 (Non *Bt.*) × MNH-456 (Bt)
	MNH-1020	96016 × MNH-456
	MNH-1026	C-26 (MNH-6070 × MNH-786) × FH-207
	MNH-886	(FH-207 × MNH-770) × Bollgard-I
CRI, Khanpur	RH-647	RH-500 × FH-113
	RH-662	319/08 × FH-113
	RH-668	VH-259 × RH-620
CRS, Sahiwal	SLH-06	SLH-334 × Neelum-121
	SLH-8	SLS-1 × FH-142
	SLH-19	SLH-336 × FH-114
CRS, Vehari	VH-327	VH-289 × VH-291 (*Bt.*)
	VH-Gulzar	VH-281 × VH-211 (*Bt.*)
	VH-189	VH-319 (*Bt.*) × FH-142 (*Bt.*)
	VH-383	VH-211 (*Bt.*) × VH-326
CRS, Bahawalpur	BH-178	(BH-162 × MNH-6070) × Neelum-121
	BH-201	(BH-172 × BH-126) × Neelum-121
	BH-221	(BH-160 × BH-176) × BH-121
CRS, Faisalabad	FH-142	FH-114 × FH-207
	FH-Lalazar	FH-207 × NuCot-N-33B (Bollgard-I)
	FH-152	FH-207 × FH-113
	FH-326	FH-942 × FH-114
	FH-490	FH-113 × FH-2006
Islamia University Bahawalpur	IUB-13	IUB-09 × MNH-789
Bandesha Seed Corporation	BS-15	IB 2009 × MNH-786
NIAB, Faisalabad	NIAB-878	B-111 × NIAB-Kiran

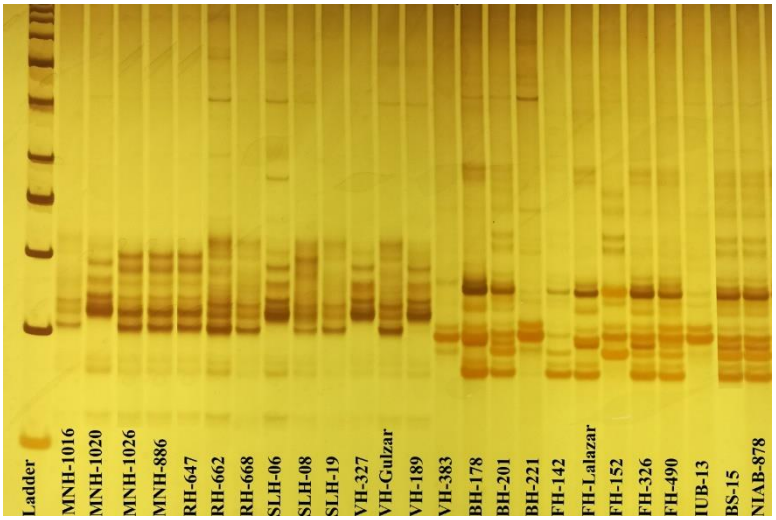

Fig. 1: The amplification product of BNL-0137. The most informative SSR marker with 19 number of alleles among which 16 are polymorphic

them to different heterotic groups for the choice of parental genotypes for hybridization-based breeding programs (Noormohammadi *et al.* 2018; Ul-Allah *et al.* 2019).

In past different types of molecular markers *i.e*., RFLPs, RAPDs, AFLPs, ISSRs, and SSR were used for DNA fingerprinting and genetic diversity studies in cotton (Becelaere *et al.* 2005; Khan *et al.* 2010; Badigannavar *et al.* 2012; Noormohammadi *et al.* 2013). However, the present study evidenced that SSR markers are still an effective tool to differentiate cotton cultivars due to high polymorphism, ease of use, and high reproducibility. However, to exploit genetic variation we have to use a very large set of SSR markers which is an indication of a narrow genetic base in the cotton germplasm (Fig. 2).

Unlike most of the previous studies (Zhang *et al.* 2013; Noormohammadi *et al.* 2018), not all the cotton varieties produced unique allelic patterns as six varieties in the present study did not amplify unique bands. Some informative SSR markers showing a high level of polymorphism are BNL0137, BNL-228, BNL0387 TMB0471, JEPSR220 BNL0140, CIR0251, BNL2616, JESPR222, BNL3590, and BNL3977.

Table 2: List of SSR markers used along with Polymorphism information, Number of Alleles (NOA) Polymorphic Alleles (PA), Polymorphic Information Contents (PIC) and annealing temperature (TA)

Sr. No.	Name	Polymorphism	NOA	PA	PIC	Sr. No.	Name	Polymorphism	NOA	PA	PIC
1.	BNL0113	Polymorphic	5	5	0.78	150.	CIR0099	Monomorphic			
2.	BNL0116	Polymorphic	4	4	0.66	151.	CIR0133	Polymorphic	8	6	0.86
3.	BNL0117	Polymorphic	8	7	0.87	152.	CIR0135	Not Amplified			
4.	BNL0118	Polymorphic	3	3	0.61	153.	CIR0143	Polymorphic	5	5	0.80
5.	BNL0119	Polymorphic	14	13	0.89	154.	CIR0169	Monomorphic			
6.	BNL0128	Polymorphic	5	5	0.78	155.	CIR0180	Monomorphic			
7.	BNL0134	Monomorphic				156.	CIR0181	Monomorphic			
8.	BNL0135	Polymorphic	4	2	0.75	157.	CIR0203	Polymorphic	7	5	0.70
9.	BNL0137	Polymorphic	19	16	0.93	158.	CIR0208	Polymorphic	2	2	0.50
10.	BNL0140	Polymorphic	14	14	0.91	159.	CIR0210	Polymorphic	2	1	0.50
11.	BNL0148	Polymorphic	6	6	0.83	160.	CIR0221	Polymorphic	3	1	0.67
12.	BNL0150	Polymorphic	10	8	0.90	161.	CIR0224	Polymorphic	4	3	0.75
13.	BNL0153	Polymorphic	7	7	0.85	162.	CIR0246	Polymorphic	6	6	0.71
14.	BNL0162A	Not Amplified				163.	CIR0251	Polymorphic	14	14	0.92
15.	BNL0174	Polymorphic	10	6	0.89	164.	CIR0272	Polymorphic	4	4	0.75
16.	BNL0193	Polymorphic	5	5	0.80	165.	CIR0294	Polymorphic	4	4	0.75
17.	BNL0197	Polymorphic	3	3	0.67	166.	CIR0307	Polymorphic	10	10	0.89
18.	BNL0206	Polymorphic	7	5	0.86	167.	CIR036	Polymorphic	8	5	0.87
19.	BNL0218	Polymorphic	10	10	0.89	168.	CIR0372	Polymorphic	7	7	0.85
20.	BNL0219	Polymorphic	3	3	0.67	169.	CIR0393	Polymorphic	6	6	0.83
21.	BNL0220	Polymorphic	3	3	0.53	170.	CIR0413	Polymorphic	3	3	0.67
22.	BNL0223	Polymorphic	10	9	0.88	171.	CIR0415	Polymorphic	5	3	0.78
23.	BNL0225	Polymorphic	12	10	0.91	172.	CIR049	Polymorphic	3	3	0.67
24.	BNL0226	Not Amplified				173.	CIR060	Polymorphic	4	1	0.75
25.	BNL0228	Polymorphic	18	18	0.92	174.	CIR062	Polymorphic	4	3	0.75
26.	BNL0234	Polymorphic	7	7	0.85	175.	CIR122	Polymorphic	7	5	0.86
27.	BNL0236	Polymorphic	4	4	0.75	176.	CM14	Not Amplified			
28.	BNL0237	Polymorphic	4	3	0.72	177.	CM17	Not Amplified			
29.	BNL0244	Polymorphic	6	5	0.82	178.	CM32	Not Amplified			
30.	BNL0285	Polymorphic	7	6	0.86	179.	CM4	Polymorphic	6	5	0.83
31.	BNL0300	Polymorphic	4	4	0.75	180.	CM45	Polymorphic	13	4	0.92
32.	BNL0329	Polymorphic	9	8	0.85	181.	CM6	Not Amplified			
33.	BNL0341	Polymorphic	9	9	0.88	182.	CM60	Polymorphic	7	3	0.86
34.	BNL0343	Polymorphic	4	4	0.74	183.	CM63	Not Amplified			
35.	BNL0347	Polymorphic	2	2	0.50	184.	CM66	Monomorphic			
36.	BNL0358	Polymorphic	5	1	0.80	185.	CM67	Polymorphic	7	7	0.76
37.	BNL0379	Not Amplified				186.	CM68	Not Amplified			
38.	BNL0386	Polymorphic	10	10	0.88	187.	CM7	Not Amplified			
39.	BNL0387	Polymorphic	17	17	0.93	188.	CM8	Not Amplified			
40.	BNL0390	Polymorphic	6	6	0.82	189.	DPL0035	Polymorphic	9	9	0.89
41.	BNL0391	Polymorphic	5	5	0.80	190.	DPL0041	Polymorphic	6	6	0.83
42.	BNL0448	Polymorphic	9	3	0.87	191.	DPL0058	Polymorphic	2	1	0.50
43.	BNL0530	Not Amplified				192.	DPL0079	Polymorphic	4	3	0.67
44.	BNL0584	Polymorphic	3	3	0.66	193.	DPL0133	Polymorphic	5	4	0.79
45.	BNL0597	Polymorphic	4	1	0.75	194.	DPL0149	Polymorphic	5	5	0.70
46.	BNL0686	Monomorphic				195.	DPL0156	Polymorphic	2	2	0.08
47.	BNL0827	Polymorphic	3	2	0.67	196.	DPL0163	Polymorphic	2	2	0.50
48.	BNL0829	Polymorphic	6	6	0.83	197.	DPL0264	Polymorphic	6	4	0.83
49.	BNL0830	Polymorphic	2	2	0.27	198.	DPL0273	Polymorphic	2	2	0.50
50.	BNL0834	Polymorphic	8	8	0.88	199.	DPL0348	Monomorphic			
51.	BNL0891	Polymorphic	5	5	0.80	200.	DPL0385	Polymorphic	4	4	0.75
52.	BNL0946	Polymorphic	5	5	0.76	201.	DPL0443	Monomorphic			
53.	BNL1017	Monomorphic				202.	DPL0489	Monomorphic			
54.	BNL1161	Polymorphic	8	4	0.87	203.	DPL0528	Polymorphic	4	3	0.75
55.	BNL1253	Polymorphic	5	5	0.77	204.	DPL0534	Polymorphic	4	3	0.75
56.	BNL1317	Polymorphic	8	8	0.84	205.	DPL0542	Polymorphic	7	4	0.82
57.	BNL1403	Polymorphic	3	3	0.67	206.	HAU0119	Polymorphic	7	3	0.82
58.	BNL1417	Polymorphic	7	7	0.84	207.	JESPR0102	Monomorphic			
59.	BNL1418	Monomorphic	4	0	0.75	208.	JESPR0135	Polymorphic	9	9	0.89
60.	BNL1441	Polymorphic	5	5	0.80	209.	JESPR0232	Polymorphic	9	7	0.89
61.	BNL1531	Polymorphic	6	3	0.83	210.	JESPR0240	Monomorphic			
62.	BNL1592	Polymorphic	2	2	0.50	211.	JESPR1	Polymorphic	4	3	0.70
63.	BNL1597	Polymorphic	9	7	0.88	212.	JESPR100	Polymorphic	4	4	0.72
64.	BNL1605	Not Amplified				213.	JESPR101	Not Amplified			
65.	BNL1667	Polymorphic	5	4	0.79	214.	JESPR103	Polymorphic	8	8	0.87
66.	BNL1681	Not Amplified				215.	JESPR108	Polymorphic	3	3	0.63
67.	BNL1688	Polymorphic	6	1	0.83	216.	JESPR114	Polymorphic	11	11	0.89
68.	BNL1694	Polymorphic	5	3	0.80	217.	JESPR134	Polymorphic	10	9	0.88
69.	BNL2443	Monomorphic				218.	JESPR153	Polymorphic	11	11	0.88
70.	BNL2448	Polymorphic	6	5	0.83	219.	JESPR156	Polymorphic	4	4	0.71
71.	BNL2527	Polymorphic	11	11	0.91	220.	JESPR160	Polymorphic	3	2	0.64
72.	BNL2544	Polymorphic	5	5	0.80	221.	JESPR173	Polymorphic	6	6	0.83
73.	BNL2564	Polymorphic	4	4	0.75	222.	JESPR178	Polymorphic	5	5	0.80
74.	BNL2570	Polymorphic	2	2	0.50	223.	JESPR185	Polymorphic	5	5	0.79
75.	BNL2572	Polymorphic	4	4	0.75	224.	JESPR186	Not Amplified			

Table 2: Continue

Table 2: Continue

76.	BNL2590	Polymorphic	9	6	0.88	225.	JESPR194	Polymorphic	8	8	0.87
77.	BNL2597	Polymorphic	5	3	0.80	226.	JESPR200	NA			
78.	BNL2599	Polymorphic	3	3	0.67	227.	JESPR202	NA			
79.	BNL2616	Polymorphic	15	14	0.91	228.	JESPR205	Polymorphic	6	4	0.82
80.	BNL2632	Polymorphic	12	12	0.88	229.	JESPR209	Polymorphic	2	2	0.50
81.	BNL2634	Polymorphic	12	10	0.91	230.	JESPR215	Polymorphic	13	13	0.91
82.	BNL2652	Polymorphic	4	4	0.73	231.	JESPR218	Monomorphic			
83.	BNL2681	Polymorphic	4	2	0.75	232.	JESPR220	Polymorphic	15	15	0.93
84.	BNL2700	Polymorphic	10	9	0.89	233.	JESPR222	Polymorphic	14	14	0.93
85.	BNL2750	Polymorphic	2	1	0.50	234.	JESPR227	Polymorphic	6	6	0.79
86.	BNL2762	Polymorphic	6	6	0.83	235.	JESPR229	Monomorphic			
87.	BNL2772	Polymorphic	5	4	0.80	236.	JESPR232	Polymorphic	7	4	0.83
88.	BNL2827	Monomorphic				237.	JESPR236	Polymorphic	7	4	0.83
89.	BNL2835	Polymorphic	11	6	0.91	238.	JESPR242	Polymorphic	6	6	0.82
90.	BNL2882	Polymorphic	9	5	0.89	239.	JESPR244	Monomorphic			
91.	BNL2986	Monomorphic				240.	JESPR246	Polymorphic	11	11	0.90
92.	BNL3029	Polymorphic	3	3	0.67	241.	JESPR250	Polymorphic	8	8	0.78
93.	BNL3034	Polymorphic	4	2	0.74	242.	JESPR270	Polymorphic	7	3	0.86
94.	BNL3071	Not Amplified				243.	JESPR272	Not Amplified			
95.	BNL3090	Polymorphic	6	3	0.83	244.	JESPR291	Monomorphic			
96.	BNL3103	Polymorphic	2	2	0.50	245.	JESPR292	Polymorphic	3	2	0.49
97.	BNL3140	Polymorphic	2	2	0.50	246.	JESPR296	Polymorphic	4	4	0.74
98.	BNL3147	Polymorphic	4	4	0.75	247.	JESPR310	Polymorphic	6	5	0.83
99.	BNL3255	Polymorphic	9	4	0.88	248.	JESPR42	Polymorphic	11	7	0.90
100.	BNL3279	Polymorphic	7	7	0.85	249.	JESPR80	Not Amplified			
101.	BNL3319	Polymorphic	5	5	0.78	250.	JESPR84	Polymorphic	8	5	0.87
102.	BNL3324	Polymorphic	3	3	0.67	251.	JESPR85	Polymorphic	2	2	0.50
103.	BNL3345	Polymorphic	5	4	0.63	252.	JESPR94	Polymorphic	3	3	0.65
104.	BNL3379	Polymorphic	7	3	0.86	253.	JESPR95	Polymorphic	7	7	0.84
105.	BNL3383	Polymorphic	7	7	0.85	254.	MGHES11a	Polymorphic	6	6	0.83
106.	BNL3408	Polymorphic	10	10	0.89	255.	MGHES11b	Polymorphic	4	4	0.73
107.	BNL3414	Polymorphic	5	3	0.80	256.	MGHES18	Polymorphic	3	3	0.65
108.	BNL3432	Polymorphic	5	5	0.80	257.	MGHES24	Polymorphic	11	11	0.91
109.	BNL3449	Polymorphic	8	8	0.86	258.	MGHES30a	Monomorphic			
110.	BNL3452	Polymorphic	5	5	0.80	259.	MGHES32	Not Amplified			
111.	BNL3523	Polymorphic	7	2	0.85	260.	MGHES40	Polymorphic	7	7	0.85
112.	BNL3556	Monomorphic				261.	MGHES41	Polymorphic	9	8	0.88
113.	BNL3558	Polymorphic	3	3	0.67	262.	MGHES44	Polymorphic	14	13	0.93
114.	BNL3563	Polymorphic	7	3	0.86	263.	MGHES46	Not Amplified			
115.	BNL3582	Polymorphic	3	2	0.66	264.	MGHES48	Polymorphic	13	13	0.92
116.	BNL3590	Polymorphic	14	14	0.92	265.	MGHES59	Polymorphic	3	3	0.64
117.	BNL3592	Polymorphic	5	3	0.80	266.	MGHES6	Polymorphic	3	3	0.56
118.	BNL3599	Polymorphic	13	13	0.92	267.	MGHES70	Polymorphic	8	6	0.80
119.	BNL3601	Polymorphic	4	2	0.71	268.	MGHES71	Polymorphic	5	5	0.79
120.	BNL3646	Polymorphic	3	1	0.67	269.	MGHES73	Polymorphic	11	11	0.89
121.	BNL3649	Monomorphic				270.	MGHES75	Polymorphic	5	5	0.76
122.	BNL3661	Polymorphic	8	5	0.86	271.	MGHES76	Polymorphic	6	6	0.82
123.	BNL3799	Monomorphic				272.	MUCS0515	Polymorphic	2	1	0.50
124.	BNL3860	Polymorphic	13	8	0.92	273.	MUSB1121	Polymorphic	5	5	0.71
125.	BNL3903	Polymorphic	4	1	0.75	274.	NAU0808	Polymorphic	5	3	0.78
126.	BNL3935	Polymorphic	11	7	0.91	275.	NAU2083	Polymorphic	9	9	0.83
127.	BNL3948	Polymorphic	3	1	0.64	276.	NAU2540	Polymorphic	5	5	0.73
128.	BNL3976	Polymorphic	3	1	0.67	277.	NAU2580	Polymorphic	4	1	0.75
129.	BNL3977	Polymorphic	14	14	0.93	278.	NAU2679	Polymorphic	4	4	0.62
130.	BNL3985	Polymorphic	4	4	0.74	279.	NAU2715	Polymorphic	4	1	0.73
131.	BNL3988	Polymorphic	6	6	0.83	280.	NAU2954	Polymorphic	5	5	0.80
132.	BNL3995	Polymorphic	3	3	0.67	281.	NAU3100	Polymorphic	8	8	0.87
133.	BNL4011	Polymorphic	3	2	0.62	282.	NAU6672	Polymorphic	4	3	0.75
134.	BNL4015	Not Amplified				283.	TMB0034	Not Amplified			
135.	BNL4030	Not Amplified				284.	TMB0471	Polymorphic	16	15	0.93
136.	BNL4078	Polymorphic	3	3	0.67	285.	TMB0603	Polymorphic	4	1	0.75
137.	BNL4080	Monomorphic				286.	TMB0770	Polymorphic	5	5	0.80
138.	BNL4082	Polymorphic	6	6	0.83	287.	TMB1296	Polymorphic	6	4	0.83
139.	BNL4092	Polymorphic	7	7	0.86	288.	TMB1356	Polymorphic	6	1	0.83
140.	BNL786	Polymorphic	7	7	0.78	289.	TMB1456	Monomorphic			
141.	BNL834	Polymorphic	7	7	0.85	290.	TMB1548	Polymorphic	6	6	0.83
142.	CGR5641	Polymorphic	4	2	0.74	291.	TMB1638	Polymorphic	4	3	0.75
143.	CGR6692	Polymorphic	4	4	0.75	292.	TMB1639	Polymorphic	6	2	0.82
144.	CGR6692	Polymorphic	3	2	0.63	293.	TMB1838	Polymorphic	3	1	0.66
145.	CGR6824	Polymorphic	7	7	0.80	294.	TMB1919	Polymorphic	5	5	0.80
146.	CIR0054	Polymorphic	4	4	0.75	295.	TMB2920	Polymorphic	2	2	0.50
147.	CIR0061	Polymorphic	7	7	0.85	296.	TMB2945	Polymorphic	3	2	0.66
148.	CIR0082	Polymorphic	11	11	0.91	297.	TMH05	Monomorphic			
149.	CIR0094	Polymorphic	10	9	0.85						

Note: Annealing temperature of all primers was 55°C

Table 3: List of SSR markers that can distinguish twenty-five varieties of cotton using direct or indirect method

Genotypes	DNA Fingerprints
MNH-886	BNL0228, MGHES24
MNH-1016	BNL0123, CIR0203, NAU2679, BNL0119, MGHES75, JESPR153
MNH-1020	BNL0119, BNL0391, BNL2634, JESPR232
MNH-1026	Identifiable using pair of SSR markers (BNL2632 & BNL0123) and (BNL0341 & CIR0230)
VH-327	MGHES75, JESPR215
VH-Gulzar	BNL0134
VH-189	Identifiable using pair of SSR markers (BNL0830 & BNL0119) and (DPL0153 & BNL0134)
VH-383	Identifiable using pair of SSR markers (BNL3601 & BNL0119) and (BNL3449 & CIR0391)
FH-142	BNL0228
FH-Lalazar	Identifiable using pair of SSR markers (BNL0830 & JESPR232) and (BNL0237 & CIR0203)
FH-152	Identifiable using pair of SSR markers (BNL834 & BNL1253) and (BNL786 & BNL448)
FH-326	DPL0542, CIR0246, DPL0149
FH-490	Identifiable using pair of SSR markers (TMB2926 & BNL0123) and (BNL3988 & JESPR232)
RH-647	BNL1253, DPL0133
RH-662	BNL2616, MGHES73
RH-668	DPL0156, CIR0094, UAU0119, BNL0329
SLH-06	BNL0448
SLH-8	MGHES6, JESPR153
SLH-19	BNL0137, JESPR250
BH-178	Identifiable using pair of SSR markers (BNL1592 & BNL3529) and (BNL0329 & JESPR153)
BH-201	JESPR236
BH-221	BNL3529, BNL0220, BNL0119
NIAB-878	NAU2083, BNL2540, BNL2599, BNL0140, JESPR114
IUB-13	HAU0119, CIR0307, BNL4082, BNL0390, BNL0150, BNL0228, BNL0119, BNL0236, JESPR100
BS-15	BNL2835, MGHES24

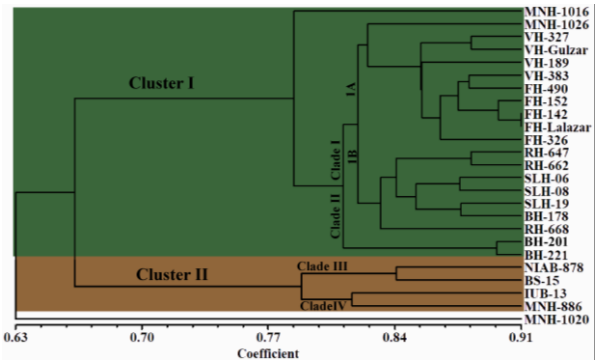

Fig. 2: Dendrogram of 25 cotton genotypes generated using data of 244 polymorphic SSR markers through SHAN similarity matrix and unweighted pair group method

These SSR markers may be used for DNA fingerprinting and genetic diversity studies in the future. Our results are in line with the previous studies (Bertini *et al.* 2006; Lacape *et al.* 2007) which also reported informative SSR markers for genotyping and genetic diversity studies.

The average alleles and polymorphic alleles per locus 6.3 and 5.3 respectively reported in our study were higher than many of previously published studies. Zhu *et al.* (2019) reported 6.02 alleles per locus in a study comprising of 557 *G. hirsutism* accessions. Javaid *et al.* (2017) reported 3.72 alleles per locus in a study of genetic diversity in 22 cotton accessions using 30 SSR markers. Similarly, Gurmessa (2019) reported 3.8 alleles per locus with 0.50 PIC value. Whereas according to our knowledge only one study of McCarty *et al.* (2018) reported a high number of alleles (7.9) per locus. This is expected because they used landraces and genetic diversity in landraces is more than the cultivated varieties. However, Average PIC value reported in our study 0.73 is highest among all the previously published reports. High number of alleles in our study and high PIC value corresponds to large set of SSR markers used in our study (Table 2).

Different studies have reported a continuous decline in cotton productivity in Pakistan for the past 03 years (Ashraf *et al.* 2018; Ali *et al.* 2019b; Rana *et al.* 2020; Jamil *et al*. 2021). Whereas some model-based future predictions are pointing out that this trend will continue for another four to five years (Ashraf *et al.* 2018). The question arises what are major factors that are hampering cotton productivity? One possible answer to this question is the lack of genetic divergence in the cultivated cotton genotypes as proved through our results. The varieties used in this study covered almost 60% of the cropped area under cotton cultivation. However, when it comes to genetics there are only two types of blood as is evident from structure analysis. About 84% of genotypes (21) have similar genetic makeup and formulate P1 (Fig. 3). The pedigree parentage dictates that five genotypes have FH-207 as a common parent. The same is the case with Neelum-121 which is used as a parent in breeding of three genotypes and many other such examples exists in Table 1.

The pressure for higher productivity in cotton farming and continuous artificial selection have narrows down the genetic base which is a major hurdle for successful cotton breeding programs (Noormohammadi *et al.* 2018). It happens when you start with a broad genetic base but if the base material (Pedigree/Parentage) is itself has narrow genetic makeup as is our case, what will be its outcome? Crops will be more prone to biotic and abiotic stresses as is happening in cotton *i.e*., Whitefly (Ahmad and Akhtar 2018), Jassids, aphids, thrips (Akhtar *et al.* 2018) and bollworms (Ahmad *et al.* 2019) heavily infest almost all

K	Reps	Mean LnP(K)	Stdev LnP(K)	Ln'(K)	|Ln''(K)|	Delta K
1	2	-18597.000000	2.404163	—	—	—
2	2	-8892.000000	1.555635	9705.000000	15887.300000	10212.743241
3	2	-15074.300000	8792.165717	-6182.300000	2002.750000	0.227788
4	2	-19253.850000	14517.538614	-4179.550000	2040.700000	0.140568
5	2	-25474.100000	3524.644462	-6220.250000	6792.450000	1.927131
6	2	-24901.900000	674.438448	572.200000	3088.500000	4.579365
7	2	-21241.200000	6023.701248	3660.700000	11133.400000	1.848266
8	2	-28713.900000	404.889343	-7472.700000	9901.750000	24.455447
9	2	-26284.850000	4165.919601	2429.050000	1047.500000	0.251445
10	2	-22808.300000	10303.252909	3476.550000	—	—

Fig. 3: Structure Analysis of Cotton varieties grown in Punjab Pakistan. Parameters: no admission model; K = 02; 10,000 Burn-in period; 100000 Rep

cotton varieties and cause almost 15–20% crop losses every year (Khan *et al.* 2016; Khanzada *et al.* 2016). Our breeding and selection efforts have narrowed down genetic base which needs to be broadened for the revival of cotton (Khanzada *et al.* 2016; Ali *et al.* 2019a).

Conclusion

DNA fingerprints were developed for twenty-five GM cotton genotypes grown in Punjab. The genetic diversity studies grouped the genotypes to two distinct groups P1 (20 genotypes), P2 (04 genotypes) whereas MNH-1020 did not follow clustering. The genetic makeup of cotton genotypes used in the study was narrow. We reported polymorphism information of 244 polymorphic SSR markers and proposed a core set of markers for future DNA fingerprinting and genetic diversity studies. Our study will provide a platform for the protection of Plant Breeders Rights and will help in registration of variety under Plant Breeders Rights Registry.

Acknowledgments

The Authors are highly thankful to Punjab Agricultural Research Board for providing funding to conduct this research work through PARB Project No. 908 entitled "DNA barcoding/fingerprinting for identification of Cotton, Wheat, Maize, Potato, Tomato and Date Palm varieties". Mr. Baber Ali Lab Assistant, for technical assistance during research work. The Cotton Research Stations and Institute across the Province for providing the plant material and technical support.

Author Contributions

SJ, RS, MZI and SUR obtained funding, SJ, RS and EY conducted research experimentation, SJ, RS conducted statistical data analysis, SJ, RS and EY drafted the manuscript, SUR and MZI critically reviewed the manuscript. SJ, RS, MZI and SUR supervised the research experimentation and all process. SJ and RS corresponded to journal for submission and review process.

Conflict of Interest

The authors declare no conflict of interest among them

Data Availability declaration

We hereby declare that data, primary or supplementary related to this article, are available with the corresponding author and will be produced on demand

References

Ahmad M, KP Akhtar (2018). Susceptibility of cotton whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae) to diverse pesticides in Pakistan. *J. Econ Entomol* 111:1834–1841

Ahmad M, B Rasool, M Ahmad, DA Russell (2019). Resistance and Synergism of Novel Insecticides in Field Populations of Cotton Bollworm *Helicoverpa armigera* (Lepidoptera: Noctuidae) in Pakistan. *J Econ Entomol* 112:859–871

Akhtar ZR, U Irshad, M Majid, Z Saeed, H Khan, AA Anjum, A Noreen, M Salman, J Khalid, M Abubakar (2018). Risk assessment of transgenic cotton against non-target whiteflies, thrips, jassids and aphids under field conditions in Pakistan. *Intl J Curr Microbiol Appl Sci* 6:11–24

Ali I, NU Khan, S Gul, SU Khan, Z Bibi, K Aslam, G Shabir, HA Haq, SA Khan, I Hussain (2019a). Genetic Diversity and Population Structure Analysis in Upland Cotton Germplasm. *Intl J Agric Biol* 22:669–676

Ali MA, J Farooq, A Batool, A Zahoor, F Azeem, A Mahmood, K Jabran (2019b). Cotton Production in Pakistan. *In: Cotton Production*, 1st Edition, p:249. Jabran K, BS Chauhan (Eds.). John Wiley & Sons Ltd., New York, USA

Allen G, M Flores-Vergara, S Krasynanski, S Kumar, W Thompson (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat Protoc* 1:2320–2325

Ashraf S, AH Sangi, ZY Hassan, M Luqman (2018). Future of cotton sector in Pakistan: A 2025 outlook. *Pak J Agric Res* 31:145–150

Badigannavar A, GO Myers, DC Jones (2012). Molecular diversity revealed by AFLP markers in upland cotton genotypes. *J Crop Improv* 26:627–640

Becelaere GV, EL Lubbers, AH Paterson, PW Chee (2005). Pedigree-vs. DNA marker-based genetic similarity estimates in cotton. *Crop Sci* 45:2281–2287

Bertini CH, I Schuster, T Sediyama, EGD Barros, MA Moreira (2006). Characterization and genetic diversity analysis of cotton cultivars using microsatellites. *Genet Mol Biol* 29:321–329

Caetano-Anolles G (1997). Resolving DNA amplification products using polyacrylamide gel electrophoresis and silver staining. *In: Fingerprinting Methods Based on Arbitrarily Primed PCR*, pp:119–134. Springer, The Netherlands

Evanno G, S Regnaut, J Goudet (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol Ecol* 14:2611–2620

Gurmessa D (2019). Genetic diversity study of improved cotton (*G. hirsutum* L.) varieties in Ethiopia using simple sequence repeats markers. *J Biotechnol* 7:6–14

Iqbal M, S Ul-Allah, M Naeem, M Ijaz, A Sattar, A Sher (2017). Response of cotton genotypes to water and heat stress: From field to genes. *Euphytica* 213:131

Jamil S, R Shahzad, SU Rahman, MZ Iqbal, M Yaseen, S Ahmad, R Fatima (2021). The level of Cry1Ac endotoxin and its efficacy against *H. armigera* in *Bt.* cotton at large scale in Pakistan. *GM Crops Food* 12:1–7

Jamil S, R Shahzad, S Kanwal, E Yasmeen, SU Rahman, MZ Iqbal (2020). DNA Fingerprinting and Population Structure of Date Palm Varieties Grown in Punjab Pakistan using Simple Sequence Repeat Markers. *Intl J Agric Biol* 23:943–950

Jans Y, WV Bloh, S Schaphoff, C Müller (2020). Global cotton production under climate change–Implications for yield and water consumption. *In: Hydrology an Earth System Sciences Discussions*, pp:1–27. https://doi.org/10.5194/hess-2019-595

Javaid A, F Awan, F Azhar, I Khan (2017). Assessment of allelic diversity among drought-resistant cotton genotypes using microsatellite markers. *Genet Mol Res* 16; Article gmr16029664

Kalia RK, MK Rai, S Kalia, R Singh, A Dhawan (2011). Microsatellite markers: An overview of the recent progress in plants. *Euphytica* 177:309–334

Khan AA, I Ashraf, G Hassan, S Ashraf (2016). On Farm Analysis of Cotton Growers Handicaps: Evidence from Cotton Belt of Pakistan. *Intl J Agric Ext* 4:79–85

Khan AI, FS Awan, B Sadia, RM Rana, IA Khan (2010). Genetic diversity studies among coloured cotton genotypes by using RAPD markers. *Pak J Bot* 42:71–77

Khanzada MS, TS Syed, S Rani, GHA Khanzada, M Salman, S Anwar, M Sarwar, AA Perzada, S Wang, AH Abro (2016). Occurrence and abundance of thrips, whitefly and their natural enemy, Geocoris Spp. on cotton crop at various localities of Sindh, Pakistan. *J Entomol Zool Stud* 4:509–515

Király I, R Redeczki, É Erdélyi, T Magdolna (2012). Morphological and molecular (SSR) analysis of old apple cultivars. *Not Bot Hortic Agrobot Cluj-Nap* 40:269–275

Lacape JM, D Dessauw, M Rajab, JL Noyer, B Hau (2007). Microsatellite diversity in tetraploid *Gossypium* germplasm: Assembling a highly informative genotyping set of cotton SSRs. *Mol Breed* 19:45–58

Lalwani S (2020). Pakistan in 2019: Navigating Major Power Relations amid Economic Crisis. *Asian Survey* 60:177–188

McCarty JC, DD Deng, JN Jenkins, L Geng (2018). Genetic diversity of day-neutral converted landrace *Gossypium hirsutum* L. accessions. *Euphytica* 214:1–14

Mumtaz AS, M Naveed, ZK Shinwari (2010). Assessment of genetic diversity and germination pattern in selected cotton genotypes of Pakistan. *Pak J Bot* 42:3949–3956

Nguyen TB, M Giband, P Brottier, AM Risterucci, JM Lacape (2004). Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. *Theor Appl Genet* 109:167–175

Noormohammadi Z, N Ibrahim-Khalili, M Sheidai, O Alishah (2018). Genetic fingerprinting of diploid and tetraploid cotton cultivars by retrotransposon-based markers. *Nucleus* 61:137–143

Noormohammadi Z, YF Hasheminejad-Ahangarani, M Sheidai, S Ghasemzadeh-Baraki, O Alishah (2013). Genetic diversity analysis in Opal cotton hybrids based on SSR, ISSR, and RAPD markers. Genetics and molecular research: *Genet Mol Res* 12:256–269

Pourabed E, MR Jazayeri Noushabadi, SH Jamali, N Moheb Alipour, A Zareyan, L Sadeghi (2015). Identification and DUS testing of rice varieties through microsatellite markers. *Intl J Plant Genomics* 2015; Article 965073

Pritchard JK, M Stephens, P Donnelly (2000). Inference of population structure using multilocus genotype data. *Genetics* 155:945–959

Rana AW, A Ejaz, SH Shikoh (2020). Cotton crop: A situational analysis of Pakistan. *In: Prepared as part of the technical assistance to ministry of national food security and research*. International Food Policy Research Institute, Government of Pakistan.

Rehman A, L Jingdong, AA Chandio, I Hussain, SA Wagan, QUA Memon (2019). Economic perspectives of cotton crop in Pakistan: A time series analysis (1970–2015)(Part 1). *J Saud Soc Agric Sci* 18:49–54

Rohlf F (1998). NTSYS-pc version 2.0. *Numerical Taxonomy and Multivariate Analysis Syste*m, pp:1–43. Exeter software, Setauket, New York

Santhy V, M Meshram, H Santosh, K Kranthi (2019). Molecular diversity analysis and DNA fingerprinting of cotton varieties of India. *Ind J Genet Plant Breed* 79:719–725

Shah ZH, M Munir, AM Kazi, T Mujtaba, Z Ahmed (2009). Molecular markers based identification of diversity for drought tolerance in bread wheat varieties and synthetic hexaploids. *Curr Issues Mol Biol* 11:101–110

Sheidai M, F Afshar, M Keshavarzi, SM Talebi, Z Noormohammadi, T Shafaf (2014). Genetic diversity and genome size variability in *Linum austriacum* (Lineaceae) populations. *Biochem Syst Ecol* 57:20–26

Shuli F, AH Jarwar, X Wang, L Wang, Q Ma (2018). Overview of the cotton in Pakistan and its future prospects. *Pak J Agric Res* 31:396–408

Singh S (2017). "White Gold" for Whom? A Study of Institutional Aspects of Work and Wages in Cotton GPNs in India. *In: Critical Perspectives on Work and Employment in Globalizing India*, pp:15–36. Springer, The Netherlands

Townsend T (2020). World natural fibre production and employment. *In: Handbook of Natural Fibres*, pp: 15–36. Elsevier, Amsterdam, The Netherlands

Ul-Allah S, S Ahmad, M Iqbal, M Naeem, M Ijaz, M Qadir, ZH Ahmad, HG Nabi (2019). Creation of new genetic diversity in cotton germplasm through chemically induced mutation. *Intl J Agric Biol* 22:51–56

Ullah I, A Iram, M Iqbal, M Nawaz, S Hasni, S Jamil (2012). Genetic diversity analysis of Bt cotton genotypes in Pakistan using simple sequence repeat markers. *Genet Mol Res* 11:597–605

UPOV-BMT B (2002). *36/10 Progress report of the 36th session of the technical committee, the technical working parties and working group on biochemical and molecular techniques and DNA-profiling in particular*. Upov-Bmt, Geneva, Switzerland

Yu J, S Jung, CH Cheng, SP Ficklin, T Lee, P Zheng, D Jones, RG Percy, D Main (2014). CottonGen: A genomics, genetics and breeding database for cotton research. *Nucl Acids Res* 42:1229–1236

Zhang Y, M Kuang, W Yang, H Xu, D Zhou, Y Wang, X Feng, C Su, F Wang (2013). Construction of a primary DNA fingerprint database for cotton cultivars. *Genet Mol Res* 12:1897–1906

Zhu L, P Tyagi, B Kaur, V Kuraparthy (2019). Genetic diversity and population structure in elite US and race stock accessions of upland cotton (*Gossypium hirsutum*). *J Cotton Sci* 23:38–47